



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : <b>C12N 9/02, 15/82</b>		A1	(11) International Publication Number: <b>WO 97/32011</b> (43) International Publication Date: <b>4 September 1997 (04.09.97)</b>
(21) International Application Number: <b>PCT/US97/03313</b>		(81) Designated States: AU, BA, BB, BG, BR, BY, CA, CN, CU, CZ, FI, GE, GH, HU, JP, KG, KR, KZ, LC, LK, LV, MD, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, TI, UA, US, UZ, VN, YU, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: <b>27 February 1997 (27.02.97)</b>			
(30) Priority Data: 60/012,705 28 February 1996 (28.02.96) US 60/013,612 28 February 1996 (28.02.96) US 60/020,003 21 June 1996 (21.06.96) US			
(71) Applicant ( <i>for all designated States except US</i> ): <b>NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH)</b>			
(72) Inventors; and			
(75) Inventors/Applicants ( <i>for US only</i> ): <b>VOLRATH, Sandra, L. [US/US]; 4225 Pine Oak Drive, Durham, NC 27707 (US). JOHNSON, Marie, A. [US/US]; 408 Heather Drive, Raleigh, NC 27606 (US). POTTER, Sharon, L. [US/US]; 3837 Whispering Branch Road, Raleigh, NC 27613 (US). WARD, Eric, R. [US/US]; 3003 Montgomery Street, Durham, NC 27705 (US). HEIFETZ, Peter, B. [US/US]; 3916 Sturbridge Drive, Durham, NC 27713 (US).</b>			
(74) Agent: <b>MEIGS, J., Timothy, 520 White Plains Road, Tarrytown, NY 10591-9005 (US)</b>			
(54) Title: <b>DNA MOLECULES ENCODING PLANT PROTOPORPHYRINOGEN OXIDASE AND INHIBITOR-RESISTANT MUTANTS THEREOF</b>			
(57) Abstract			
<p>The present invention provides novel DNA sequences coding for plant protoporphyrinogen oxidase (protox) enzymes from soybean, wheat, cotton, sugar beet, grape, rice and sorghum. In addition, the present invention teaches modified forms of protox enzymes that are resistance to protox inhibitors via mutation of the native protox gene to a resistant form or they may be transformed with a gene encoding an inhibitor-resistant form of a plant protox enzyme.</p>			

Soybean Protox-1, in the pBluescript SK vector, was deposited December 15, 1995 as pWDC-12 (NRRL #B-21516).

An alignment of the predicted amino acid sequences of the respective proteins encoded by the sequences shown in SEQ ID NOS: 2, 6, 10, 12, 15, 17, 19, 21, 23 and are set forth in Table 1. An alignment of the predicted amino acid sequences of the respective proteins encoded by the sequences shown in SEQ ID NOS: 4 and 8 are set forth in Table 2.

TABLE 1

Comparison of Protox-1 Amino Acid Sequences from Arabidopsis ("Arabpt-1"; SEQ ID NO:2), Maize ("Mzpt-1"; SEQ ID NO:6), Wheat ("Wtpt-1"; SEQ ID NO:10), Soybean ("Soybeanpt-1"; SEQ ID NO:12), Cotton ("Cottonpt-1"; SEQ ID NO:16), Sugar beet ("Sugpt-1"; SEQ ID NO:18), Rape ("Rapept-1"; SEQ ID NO:20), Rice ("Ricept-1"; SEQ ID NO:22), and Sorghum ("Sorghumpt-1"; SEQ ID NO:24)

Alignment is performed using the PileUp program (GCG package, University of Wisconsin, Madison, WI). Positions that may be modified according to the teachings herein to confer or enhance inhibitor resistance are shown in bold type.

	1		50
Rapept-1	.....	M DLSLLRP..	QPFILSPFSNP FPRSRPYKPL
Arabpt-1	.....	M ELSLLRPTT	QSILLPSFSKP NLRLNVYKPL
Sorghumpt-1	.....	.....	.....
Mzpt-1	.....	.....	.....
Wtpt-1	.....	.....	M ATATVAAAASP LRGRVTGRPH
Ricept-1	.....	.....	.....
Cottonpt-1	.....	M TAL IDLSLLRSSP	S VSPF S I P H I Q H P P R F R K P F
Soybeanpt-1	.....	M V S V F N E I L F P P	N Q T I L R P S L H S P T S F F T S P T R K F P R S R P N P
Sugpt-1	M K S M A L S N C I	P Q T Q C M P L R S	S G H Y R Q N C I M L S I P C S L I G R R G Y Y S H K R R
	51		100
Rapept-1	N L R C S V S G G S	W G S S T I E G G	G X G X X G
Arabpt-1	L L R C S V A G G P	T V G S S K I E G G	G G T . T I T T D C
Sorghumpt-1	.....	.....	V I V G G G I S G L C T A Q A L A T K H
Mzpt-1	.....	.....	A D C V V V G G G I S G L C T A Q A L A T R H



Wtpt-1	RVRPRCATAS SATETPAAPG VRL...SAEC VIVGAGISGL CTAQALATRY	
Ricept-1	.....	
Cottonpt-1	KLRCSLAEGP TISSSKILDGG ESS...IADC VIVGGGIISGL CIAQALATKH	
Soybeanpt1	ILRCSIAEES TASPPKTR.. DSA...PVDC VVGGGVSGL CIAQALATKH	
Sugcpt-1	MSMSCSTSSG SKSAVKEAGS GSGAGGLDC VIVGGGIISGL CIAQALCTKH GxGxxG	←
 101		150
Rapept-1	PDA..AKNVM VTEAKDRVGG NIIT..REQ GFLWEEGPNS FQPSDPMUIM	
Arabpt-1	PDA..APNLI VTEAKDRVGG NIIT..REEN GFLWEEGPNS FQPSDPMUIM	
Sorghumpt-1	..... STVERPEE GYLWEEGPNS FQPSDPVLSM	
Mzpt-1	..G..VGDL VTEARARPGG NITTVERPEE GYLWEEGPNS FQPSDPVLM	
Wtpt-1	..G..VSDLL VTEARDRPGG NITTVERPDE GYLWEEGPNS FQPSDPVLM	
Ricept-1	.....	
Cottonpt-1	RDV..ASNVI VTEARDRPGG NITTVER..D GYLWEEGPNS FQPSDPMLIM	
Soybeanpt1	..A..NANVV VTEARDRPGG NITIMER..D GYLWEEGPNS FQPSDPMLIM	
Sugcpt-1	SSSSLSPNFI VTEAKDRVGG NIVTE..AD GYLWEEGPNS FQPSDAVLIM	
 151		200
Rapept-1	VVDSGLKDDL VLGDPTAPRF VLWNGKLRPV PSKLTDLPFF DLMSIGGKIR	
Arabpt-1	VVDSGLKDDL VLGDPTAPRF VLWNGKLRPV PSKLTDLPFF DLMSIGGKIR	
Sorghumpt-1	AVDGLKDDL VFGDPNAPRF VLWEGKLRPV PSKPADLPFF DLMSIPGCLR	
Mzpt-1	AVDGLKDDL VFGDPNAPRF VLWEGKLRPV PSKPADLPFF DLMSIPGCLR	
Wtpt-1	AVDGLKDDL VFGDPNAPRF VLWEGKLRPV PSKPADLPFF SLMSIPGCLR	
Ricept-1	.....	
Cottonpt-1	AVDGLKDDL VLGDPNAPRF VLWEGKLRPV PSKPADLPFF DLMSIAGKLR	
Soybeanpt1	VVDSGLKDEL VLGDOPDAPRF VLWNRKLRPV PGKLTDLPFF DLMSIGGKIR	
Sugcpt-1	AVDGLKDEL VLGDPNAPRF VLWNDKLRPV PSSLTDLPFF DLMTIPGKIR	
 201		250
Rapept-1	AGFGAIGIRP SPPGREESVE EFVRRNLGDE VFERLIEPFC SGVYAGIDPAK	
Arabpt-1	AGFGALGIRP SPPGREESVE EFVRRNLGDE VFERLIEPFC SGVYAGIDPSK	
Sorghumpt-1	AGLGALGIRP PAPGREESVE EFVRRNLGAE VFERLIEPFC SGVYAGIDPSK	
Mzpt-1	AGLGALGIRP PPPGREESVE EFVRRNLGAE VFERLIEPFC SGVYAGIDPSK	
Wtpt-1	AGLGALGIRP PPPGREESVE EFVRRNLGAE VFERLIEPFC SGVYAGIDPSK	
Ricept-1	.....	
Cottonpt-1	AGFGAIGIRP PPPGYEEESVE EFVRRNLGAE VFERFIEPFC SGVYAGIDPSK	

Soybeanpt1 AGFGALGIRP PPPGHEESVE EFVRRNLGDE VFERLIEPPC SGVYAGDPSK  
 Sugpt-1 AALGALGFRP SPPPHEESVE HFVRRNLGDE VFERLIEPPC SGVYAGDPAK

251

300

Rapept-1 LSMKAAGFKV WKLEENGSI IGGIFKAIQD KNKAPKTTRD PRLPKPKGQT  
 Arabpt-1 LSMKAAGFKV WKLEQNGSI IGGIFKAIQE RKNAPKAERD PRLPKPKGQT  
 Sorghumpt-1 LSMKAAGFKV WRLEETAGGSI IGGITIKTIQE RGKNPKPPRD PRLPKPKGQT  
 Mzpt-1 LSMKAAGFKV WRLEETGGSI IGGITIKTIQE RSKNPKPPRD ARLPKPKGQT  
 Wtpt-1 LSMKAAGFKV WRLEETIGGSI IGGITIKAIQD KGKNPKPPRD PRLPAPKGQT  
 Ricept-1 RALKAAFGRV WRLEDIGGSI IGGITIKTIQE RGKNPKPPRD PRLPTPKGQT  
 Cottonpt-1 LSMKAAGFGRV WKLEETIGGSI IGGITFKTIQE RNKTPKPPRD PRLPKPKGQT  
 Soybeanpt1 LSMKAAGFKV WKLEKN3GSI IGGIFKAIQE RNGASKPPRD PRLPKPKGQT  
 Sugpt-1 LSMKAAGFKV WKLEQKGGSI IGGITLKAIQE RGSNPKPPRD QRLPKPKGQT

301

350

Rapept-1 VGSFRKGLIM LPEAISARLG DKVKVSWKLS SITKLASGEY SLTYETPEGI  
 Arabpt-1 VGSFRKGLRM LPEAISARLG SKVLSWKLS GITKLESGGY NLTYETPDGL  
 Sorghumpt-1 VASFRKGLAM LPNAITSSLG SKVLSWKLT SMKSDGKGY VLEYETPEGV  
 Mzpt-1 VASFRKGLAM LPNAITSSLG SKVLSWKLT SITKSDGKGY VLEYETPEGV  
 Wtpt-1 VASFRKGLAM LPNAIASRLG SKVLSWKLT SMKADNQGY VLGYETPEGI  
 Ricept-1 VASFRKGLIM LPDAITSRLG SKVLSWKLT SITKSDNKGY ALVYETPEGV  
 Cottonpt-1 VGSFRKGLIM LPEAISANSLG SNVLSWKLS SITKLNGGY NLTFETPEGM  
 Soybeanpt1 VGSFRKGLIM LPDAISARLG NKVLSWKLS SISKLDNSEY SLTYETPEGV  
 Sugpt-1 VGSFRKGLVM LPDAISARLG SRVLSWTLs SIVKSINGEY SLTYDTIPDGL

351

400

Rapept-1 VTQSKSVVM TVPSHVASSL LRPLSDSAAE ALSKLYYPPV AAVSISYAKE  
 Arabpt-1 VSVQSKSWM TVPSHVASGL LRPLSESAAN ALSKLYYPPV AAVSISYPKE  
 Sorghumpt-1 VLVQAKSVIM TIPSYVASDI LRPLSGDAAD VLSRFYYPPV AAVTVSYPKE  
 Mzpt-1 VSVQAKSVIM TIPSYVASNI LRPLSSDAAD ALSRFYYPPV AAVTVSYPKE  
 Wtpt-1 VSVQAKSVIM TIPSYVASDI LRPLSIDAAD ALSKFYYPPV AAVTVSYPKE  
 Ricept-1 VSVQAKTVM TIPSYVASDI LRPLSSDAAD ALSIFYYPPV AAVTVSYPKE  
 Cottonpt-1 VSLQSRSSVM TIPSHVASNL LHPLSAAAAD ALSQFYYPPV ASVTVSYPKE  
 Soybeanpt1 VSLOCKTVVL TIPSYVASTL LRPLSAAAAD ALSKFYYPPV AAVSISYPKE  
 Sugpt-1 VSVRTKSVVM TVPSYVASRL LRPLSDSAAAD SLSKFYYPPV AAVSLSYPKE

	401	450
Rapept-1	AIRSECLIDG ELKGFQQLHP RTQKVEILGT IYSSSLFPNR APPGRVLLIN	
Arabpt-1	AIRTECLIDG ELKGFQQLHP RTQGVEILGT IYSSSLFPNR APPGRILLIN	
Sorghumpt-1	AIRKECLIDG ELOGFQQLHP RSQGVETLGT IYSSSLFPNR APAGRVLLIN	
Mzpt-1	AIRKECLIDG ELOGFQQLHP RSQGVETLGT IYSSSLFPNR APDGRVLLIN	
Wtpt-1	AIRKECLIDG ELOGFQQLHP RSQGVETLGT IYSSSLFPNR APAGRVLLIN	
Ricept-1	AIRKECLIDG ELOGFQQLHP RSQGVETLGT IYSSSLFPNR APAGRVLLIN	
Cottonpt-1	AIRKECLIDG ELKGFQQLHP RSQGVETLGT IYSSSLFPNR APSGRVLLIN	
Soybeanpt1	AIRSECLIDG ELKGFQQLHP RSQGVETLGT IYSSSLFPNR APPGRVLLIN	
Sugpt-1	AIRSECLING ELOGFQQLHP RSQGVETLGT IYSSSLFPGR APPGRILILS	

	451	500
Rapept-1	YIGGMINTGI LSKSEGELVE AVDRDLRKML IKPSSTDPLV LGVKLWPQAI	
Arabpt-1	YIGGSINTGI LSKSEGELVE AVDRDLRKML IKPNSTDPLK LGVRVWPQAI	
Sorghumpt-1	YIGGMINTGI VSKTESELVE AVDRDLRKML INPTAVDPLV LGVRVWPQAI	
Mzpt-1	YIGGMINTGI VSKTESELVE AVDRDLRKML INSTAVDPLV LGVRVWPQAI	
Wtpt-1	YIGGSINTGI VSKTESDLVG AVDRDLRKML INPRAADPLA LGVRVWPQAI	
Ricept-1	YIGGSINTGI VSKTESELVE AVDRDLRKML INPRAVDPLV LGVRVWPQAI	
Cottonpt-1	YIGGMINTGI LSKTEGELVE AVDRDLRKML INPAKDPLV LGVRVWPKAI	
Soybeanpt1	YIGGMINTGI LSKTIDSELVE TVDRDLRKIL INPNAQDPFV VGVRILWPQAI	
Sugpt-1	YIGGAKNPGI LNKSNDLAK TVIKDLRRML INPKDAKLPRV LGVRVWPQAI	

	501	550
Rapept-1	PQFLIGHIDL VDAAKASLSS SGHEGLFLGG NYVAGVALGR CVEGAYETAT	
Arabpt-1	PQFLVGHFDI LTDIAKSSLTS SGYEGFLGG NYVAGVALGR CVEGAYETAI	
Sorghumpt-1	PQFLVGHLDL LEAAKSALDQ GGYNGLFLGG NYVAGVALGR CIEGAYESAA	
Mzpt-1	PQFLVGHLDL LEAAKAALDR GGYDGLFLGG NYVAGVALGR CVEGAYESAS	
Wtpt-1	PQFLIGHIDL LAAAKSALGQ GGYDGLFLGG KYVAGVALGR CIEGAYESAS	
Ricept-1	PQFLIGHIDL LEAAKSALGK GGYDGLFLGG NYVAGVALGR CVEGAYESAS	
Cottonpt-1	PQFLVGHLDL LDIAKSMALRD SGFHGLFLGG NYVSGVALGR CVEGAYEVAA	
Soybeanpt1	PQFLVGHLDL LDVAKASIRN TGWEGLFLGG NYVSGVALGR CVEGAYEVAA	
Sugpt-1	PQFSIGHFDL LDAAKAALTD TGVKGLFLGG NYVSGVALGR CIEGAYESAA	

	551	563
Rapept-1	QVNDFMSRYA YK*	
Arabpt-1	EVNNFMSRYA YK*	

Sorghumpt-1	QIYDFLTLYA YK*
Mzpt-1	QISDFLTLYA YK*
Wtpt-1	QVSDFLTLYA YK*
Ricept-1	QISDYLTLKYA YK*
Cottonapt-1	EVKEFLSQYA YK*
Soybeanapt1	EVNDFLINRV YK*
Sugpt-1	EVVDFLSQYS DK*

**TABLE 2**

**Comparison of the Arabidopsis (SEQ ID NO:4) and  
Maize (SEQ ID NO:8) Protox-2 Amino Acid Sequences**

Identical residues are denoted by the vertical bar between the two sequences. Alignment is performed using the GAP program described in Devereaux et al., *Nucleic Acids Res.* 12:387-395 (1984).

Percent Similarity: 75.889 Percent Identity: 57.905

Protox-2.Pep x Mzprotox-2.Pep

1	.....MASGAVAD.HQIEAVSGKRVAV	21
	::  .. ...	
1	MLALTASASSASSHPYRHASAHTRRPRLRAVLAMAGSDDPRAAPARSAV	50
22	VGAGVSGLAAAYKLKSRLNVTVFEADGRVGGKLR SVMQNGLIWDEGANT	71
	: : . :     ..: ..   : . .: :	
51	VGAGVSGLAAAYRLRQSGVNVTVFEAADRAGGKIRTNSEGGFWDEGANT	100
72	MTEAEPEVGSLDDLGRLREKQQFPISQKKRYIVRNGVPVMLPTNPIELVT	121
	:   ... :    ..:   :    .   :  : ..: ..: .. : .. :	
101	MTEGEWEASRLIDDLGLQDKQQYPNSQHKRYIVKDGA PALIPS DPLS LMK	150
122	SSVLSTQSKFQILLEPFLWKK....KSSKVSDASAEE SVSEFFQRHFGQE	167
	.  :...:   :   .:   :.. .   :  :   .	
151	SSVLSTKSKIALFFEPFLYKKANTRNSGKVSEEHLS ESVG SF CERHFGRE	200